

大豆蛋白质含量遗传的基因效应分析*

孟祥勋¹ 王曙明¹ 杨庆凯²

(1 吉林省农科院大豆所 公主岭 136100 2 东北农业大学大豆所)

摘 要

利用两个杂交组合(组合 I:吉林 27 \times 吉林 28;组合 II:吉林 20 \times 吉林 26)的 P₁、B₁、F₁、F₂、B₂、B₃ 六个世代,在对世代间蛋白质含量方差分析显著的基础上,以世代平均值模型尺度检验,三参数和六参数联合尺度的多元回归分析检验,研究探讨了大豆子粒蛋白质含量遗传的 [d] [h] [i] [j] [l] 各种基因效应及其相对效应估值的大小。世代方差分析结果表明,不同世代间蛋白质含量存在显著差异。模型尺度检验及回归分析的联合尺度检验结果表明,2 个组合中有 1 个组合(组合 I)蛋白质含量符合三参数的加性-显性基因效应模式,而且加性效应为主,其加性效应估值近显性效应的 3 倍;六参数模型分析时,两组合加性和显性效应均达极显著;其相对平方和大小,组合 I 分别占总平方和的 47.88% 和 44.76%,组合 II 分别占 34.53% 和 43.98%,表现为加性和显性效应同等重要。组合 II 模型尺度检验中不符合加性-显性模式,六参数模式分析时,上位效应 [i] 亦显著,占总平方和的 16.22%。

关键词 大豆;蛋白质;基因效应

引 言

育种家在对作物数量性状的遗传改良中,为设计合理有效的育种方案及选择方法,通常针对目标性状的遗传变异,估计其各种遗传参数,分析决定该性状的基因作用方式及其大小。就大豆子粒蛋白质含量的基因效应,学者们采用各种试验设计统计分析方法进行了估计。如世代平均数比较(Weber 等 1950),N_D 设计分析(Leffel 和 Weiss 1958,刘显华 1990),双列杂交配合力分析(Ishige 1984, Chauhan 和 Singh 1983, Gupta 等 1982,陈恒鹤 1989),这些研究结果不尽一致。

本试验利用普通蛋白含量水平的 2 个品种,与蛋白含量较高的新近育成的 2 个高蛋

* 该作者现已调入苏州大学生物系工作, 215151

收稿日期 1999-05-20

Received on May 20, 1999

白品种,配制了两组合,分别以模型尺度,联合尺度检验和世代均数多元回归分析方法对蛋白质含量遗传的基因效应进行分析,估计各种基因效应方差的大小及其百分比,以便进一步明确大豆蛋白质含量遗传本质

材料及方法

利用蛋白质含量不同的 4 个品种配制 2 个杂交组合(表 1),并将 F_1 与亲本回交,形成 P_1 、 P_2 、 F_1 、 F_2 、 B_1 和 B_2 六个世代。

表 1 杂交组合、亲本及其蛋白质含量

Table 1 Parents and their protein content for crosses

组合 Cross	亲本 Parent	蛋白质含量 (%) Protein content	组合 Cross	亲本 Parent	蛋白质含量 (%) Protein content
I	♀ 吉林 28号	46.6	II	♀ 吉林 26号	46.4
	♂ 吉林 27号	40.3		♂ 吉林 20号	39.2

1993 年种植于吉林省农科院(公主岭)试验田。随机区组,三次重复, F_1 、 B_1 、 B_2 一个行区, P_1 、 P_2 、 F_2 为 3 行区,行长 300cm,行距 65cm 株距 15cm 成熟后每个小区收获 5-6 株,单株脱粒,分析测定蛋白质含量

采用下列方法对蛋白质含量的遗传进行了分析:

(1)以 Mather(1949)提出 A B C 模型,

$$A = 2B_1 - P_1 - F_1;$$

$$B = 2B_2 - P_2 - F_1;$$

$$C = 4E - 2F_1 - P_1 - P_2 \text{ 进行 } t \text{ 测验。}$$

(2)以 Cavill(1952)提出的联合尺度检验:

首先以各世代平均数估算加性-显性模型的 m [a] 和 [d],然后由计算出的 m [a] [d] 按加性-显性模型计算各世代平均数的期望值,最后将各世代的平均观察值与计算出的期望值进行 χ^2 检验,若 χ^2 检验期望值与观察值吻合,证明加性-显性模型适合,反之模型则不适合。

表 2 各世代期望值估计的理论系数

Table 2 Theoretical coefficients of expect estimates for 6 generations

基因效应 Gene effect	世代 Generation						
	P_1	P_2	F_1	F_2	B_1	B_2	
m	1	1	1	1	1	1	
[d] 加性效应 additive	1	-1	0	0	1/2	-1/2	
[h] 显性效应 dominance	0	0	1	1/2	1/2	1/2	
[i] 加性×加性效应 additiv×additive	1	1	0	0	1/4	1/4	
[j] 加性×显性效应 additiv×dominance	0	0	0	0	1/4	-1/4	
[l] 显性×显性效应 dominance×dominance	0	0	1	1/4	1/4	1/4	

根据上表,求解三参 m [a] [d] 正规方程:

$$XA = Y$$

X 为 m [d] [h] 的理论系数矩阵, Y 为各世代平均数矩阵, A 为遗传参数矩阵

(3) 根据 (2) 原理, 采用 Ink 和 Jones (1958) 参数模型:

$$m = 1/2P_1 + 1/2P_2 + 4F_2 - 2B_1 - 2B_2;$$

$$[d] = 1/2P_1 - 1/2P_2;$$

$$[h] = 6B_1 + 6B_2 - 8F_2 - F_1 - 3/2P_1 - 3/2P_2;$$

$$[i] = 2B_1 + 2B_2 - 4F_2;$$

$$[j] = 2B_1 - P_1 - 2B_2 + P_2;$$

$$[l] = P_1 + P_2 + 2F_1 - 4F_2 - 4B_1 - 4B_2;$$

同样, 可求出 [d] [h] [i] [j] [l] 五个参数。

结果与分析

1 世代方差分析

两个组合蛋白质含量的世代间方差估值列于表 3。由表 3 可见, 蛋白质含量两组合均表现有显著的世代差异, 说明世代不同群体的遗传结构是不同, 也是进一步分析基因效应的基础。

表 3 世代间蛋白质含量方差分析

Table 3 Variance analysis of protein content among generations

变异来源 Variation resource	DF	组合 I Cross I		组合 II Cross II	
		SS	F	SS	F
世代 Generation	5	33.441	30.6 *	51.132	119.18 *
区组 Replication	2	0.022	0.05	1.073	6.25
机误 Error	10	2.181		0.858	

2 蛋白质含量遗传的 A B C 模型检验

两组合蛋白质含量遗传的 A B C 模型检验结果列于表 4。结果看出, 两组合测验结果不同, 组合 I A B C 尺度均与零没有显著差异, 表明该组合蛋白质含量符合加性 - 显性基因效应, 而组合 II A B C 尺度与零有显著的差异, 则表明不拟合简单的加性 - 显性遗传模式, 可能还有上位作用的存在。

表 4 两组合蛋白质遗传的 A B C 尺度检验

Table 4 A B C scale test of inheritance of protein content in two crosses

组合 Cross	模型尺度 Test-scale	方差 Variance	t 值 t-value	显著性 Significance	组合 Cross	模型尺度 Test-scale	方差 Variance	t 值 t-value	显著性 Significance
I	A	0.4002	0.2048	N	II	A	0.3186	2.9016	S
	B	0.7091	0.2728	N		B	0.3290	1.2813	N
	C	1.7567	0.5849	N		C	2.6291	2.9880	S

3 蛋白质含量基因效应联合尺度检验

表 5 两组合蛋白质含量加性 - 显性 χ^2 测验

Table 5 χ^2 test of additive-dominant model for protein content in two crosses

组合 Cross	世代 Generation	平均数组成 Mean composition	观察值 (O) Observed	期望值 (E) Expected	O- E 差 O- E	总 χ^2 Total χ^2
I	P ₁	m ⁺ [d]	46.191	46.215	- 0.0242	1.0408
	P ₂	m ⁻ [d]	41.038	41.088	- 0.0506	
	F ₁	m ⁺ [h]	41.142	43.279	- 0.1374	
	F ₂	m ⁺ 1/2[h]	43.727	43.466	0.2613	
	B ₁	m ⁺ 1/2[d]+ 1/2[h]	44.752	44.747	0.0044	
	B ₂	m ⁻ 1/2[d]+ 1/2[h]	42.154	42.180	- 0.0304	
II	P ₁	m ⁺ [d]	46.162	46.009	0.1820	14.139
	P ₂	m ⁻ [d]	42.079	41.982	0.0969	
	F ₁	m ⁺ [h]	43.903	43.572	0.3312	
	F ₂	m ⁺ 1/2[h]	42.868	43.784	- 0.9159	
	B ₁	m ⁺ 1/2[d]+ 1/2[h]	44.348	44.791	- 0.4426	
	B ₂	m ⁻ 1/2[d]+ 1/2[h]	42.624	42.777	- 0.1534	

$$df=3, \chi_{(0.05)}^2=7.82$$

A B C单独检验中,每个尺度只利用了几个世代,如 A尺度仅利用了 P₁ P₂ B₁ 三个世代,未同时使用其它世代。另外在计算尺度时,又把各世代同等对待,即各个世代的平均数具有相等的标准误。而实际上,各世代方差并不相同,不一定具有相同的标准误。因此,有必要将各个世代的平均值除以相应的标准误,标准化处理后进行联合尺度测验。表 5为联合尺度检验,观察与期望值的 χ^2 测验结果。

从表 5可见,加性 - 显性模型检验结果与 A B C尺度一致。即组合 I 总 $\chi^2=1.0408$, 小于 $\chi_{0.05}^2(7.82)$ 的水平,观察值与期望值完全吻合;而组合 II 总 $\chi^2=14.1395$, 大于 $\chi_{0.05}^2$ 水平,观察值与期望值不吻合,表明仍有其它基因效应 [i]起作用。

4 蛋白质含量的基因效应估值与平方和分割

两个组合蛋白质含量的三参数和六参数基因效应估值列于表 6。从表 6可见,三参数模型中,二组合 m值接近双亲平均值,而六参模式中 m值则有所不同,其与双亲均值差距的大小取决于 [i] 的上位效应估值。组合 I 符合加性 - 显性模式, [i] 估值较小,三参数与六参数 m值相差较小;而组合 II [i] 效应较大,表明该组合蛋白质含量的遗传还包含有一定的加性×加性的上位效应。

通过世代平均值分析,求得各类基因效应平方和及其平方和占总遗传方差的比例(表 7)。

表 7结果表明,两组合蛋白质含量加性和显性效应均达极显著水平,合计分别占总遗传效应的 92.64% 和 77.51%。表明加性和显性是主要的。其中组合 I 加性和显性分别占 47.88% 和 44.76%,基本上各占总效应的近 50%,属于典型的加 - 显性遗传模式;而组合 II 除加性和显性效应各占 34.53% 和 42.98% 外,还有 16.01% 加性×加性效应,并达显

著水平,也是不可忽略。

表 6 两组合不同模式的基因效应估值

Table 6 Gene effect estimation of protein content of different models in two crosses

组合 Cross	模式 Model	基因效应 Gene effect					
		m	[d]	[h]	[i]	[j]	[l]
I	三参数 Three parameter	43.940	2.3303	-0.8049			
	六参数 Six parameter	44.721	2.5669	-2.3983	-1.0970	0.0626	0.8189
II	三参数 Three parameter	44.218	1.8480	-1.1617			
	六参数 Six parameter	41.668	2.0559	2.5653	2.4674	-0.6674	-0.3309

表 7 蛋白质含量世代平均值平方和分割

Table 7 Partition of sum square of protein generation mean content

变异来源 V. source	DF	组合I Cross I		组合II Cross II	
		平方和	基因效贡献率(%)	平方和	基因效贡献率(%)
		SS	Rate of gene action	SS	Rate of gene action
[d]	1	8.076 *	47.886	17.110 *	34.532
[h]	1	7.549 *	44.763	21.300 *	42.986
[ad]	3	1.239	7.345	11.138 *	22.480
[i]	1	0.687	4.073	7.935	16.014
[j]	1	0.039	0.002	2.146	4.332
[l]	1	0.513	3.039	1.058	2.135
世代	5	16.864	100	38.411	100

讨 论

1 基因效应与亲本表现

加性、显性、上位性各类基因效应估值是有关性状位点效应的综合估值。不同研究或不同杂交组合双亲有差异的位点或数目不同,基因效应估值就会有所不同,尤其是有关位点存在互作或作用方向不同时,不同亲本配制的组合的基因效应估值可能有相当大的差异。本研究中所采用的两个组合,父母本蛋白质含量表现虽然相近似,但其基因效应模式和效应大小都不尽相同,表明蛋白质含量的遗传方式是比较复杂的。

2 基因效应与参数模式

三参数模式就是通常所说的加性-显性模式,其中 m 值乃是无上位情况下的双亲平均值。因此,若上位效应可忽略不计,则双亲实际平均值应接近于 m 估值。如果上位效应显著,则二者间就会有差距,差距大小取决于上位效应的作用大小。进行三参数模式分析后,应对各基因效应进行显著性检验,如果离差(Dev)项不显著,说明加性-显性模式适合,那么上位效应可以忽略,[d]和[h]的作用足以说明绝大部分的遗传变异,反之若离差(Dev)显著,说明上位效应确实存在,则需要采用完全模式(六参数)进一步分析上位效应的作用。

3 蛋白质含量的基因效应和遗传改良

本文对两个杂交组合的分析表明,大豆子粒蛋白质含量的遗传,因组合不同基因效应模式有差异。并且在三参数模式中,加性效应是主要的,两组合的加性效应估值分别为显性效应的3倍和1.5倍,加性效应为主。因此说,在附合加性-显性模式的杂种后代,蛋白质含量遗传加性效应是重要的。这可能也是某些研究结果认为加性效应是主要的缘故(Weber 1950, Chuaham 1983)。在六参数模式中,加性和显性都是主要的,Leffel(1958)、Ishige(1984)、张国栋等(1989)亦都得到类似的结果。综合前人研究结果及本文结果分析,可以认为,蛋白质含量的基因效应是复杂的。因组合亲本不同,有的可能表现为简单的加性效应为主,有的则可能附合加性-显性模式,还有的组合有一定的上位效应。因此,由于基因作用的复杂性,在改良大豆蛋白质含量的育种中,虽然可大致通过亲本表现预测后代的变异及分布,但在后代处理上,适当推迟选择世代是有益的。

参 考 文 献

- [1] 刘显华, 1988, 大豆杂种第二代种子蛋白质、脂肪及其组分的配合力与遗传分析, 作物学报, 14(4): 303-307
- [2] 张国栋, 王金陵等, 1989, 大豆种间杂交主要农艺性状和蛋白质含量的遗传变异研究, 大豆科学, 8(1): 1-9
- [3] 陈恒鹤, 梁振富等, 1988, 大豆蛋白质及脂肪含量的遗传和选择效果研究。正反交效应及配合力分析, 大豆科学, 7(3): 217-225
- [4] 胡明祥等, 1984, 大豆杂种后代籽粒蛋白质含量的遗传研究, 中国农业科学, 6(6): 40-44
- [5] Chuahan V. S., and B. B. Singh. 1983, Genetic analysis of protein and oil content in soybean. Indian J. Agric. Sci. 53(8): 634-637
- [6] Gupta A. K. et al., 1982, Genetic analysis of yield and quality traits in soybean (*Glycine max* (L.) merrill). Legume Research 5(1): 49-53
- [7] Ishige T., 1984, Biometrical analysis and estimation of the number of genes for seed protein content of soybean *Glycine max* (L.) merill. Japan. Agric. Res. Quarterly 17(4): 230-235
- [8] Leffel R. C. and M. G. Weiss., 1985, Analysis of diallel crosses among ten varieties of soybean. Agron. Journal 50: 528-534
- [9] Mckendry A. L. 1985, Inheritance of seed protein and seed oil content in early maturing soybean. Can. Jour. Genet. Cytology 27(5): 603-607
- [10] Weber C. R., 1950, Inheritance and interrelation of some agronomic and chemical characters in an interspecific cross in soybeans *G. max* × *G. usuriensis*. Iowa Agric. Exp. Res. Bull. 374: 767-816

ANALYSIS OF GENETIC EFFECT OF SEED PROTEIN CONTENT IN SOYBEAN CROSSES

Meng Xiangxun¹ Wang Shuming¹ Yang Qingkai²

(1 *Soybean Institute, Jilin Academy of Agricultural Sciences* 136100

2 *Institute of Soybean, Northeast Agricultural University*)

Abstract

Gene actions, including additive [d], dominance [h], epistasis [i] of seed protein content and magnitudes of these effect were estimated with means of P, P₂, F₁, F₂, B₁, B₂ parents and generations of two crosses by Joint Scaling Test and multiple linear regression of three-parameter and six-parameter models. It was indicated that protein content of soybean seeds was significantly conditioned by additive and dominate effects in one of the two crosses, fitting an additive-dominate model, in which additive was more important, while another cross had significant additive, dominate and epistatic effects for protein content of seeds, accounting for 34.5%, 43.9% and 16.2% of the total effects respectively, fitting additive-dominate-epistatic model.

Key words Soybean; Protein; Genetic effect